DOCUMENT RESUME

ED 025 331 PS 001 492

By- Lane, Harlan; Sheppard, William

Development of the Prosodic Features of Infants' Vocalizing.

Michigan Univ., Ann Arbor. Center for Human Growth and Development.
Spons Agency- National Inst. of Child Health and Human Development, Bethesda, Md.
Report No- UM-CHGD-8
Pub Date 30 Nov 65
Note- 12p.
EDRS Price MF-$0.25 HC-$0.70
Descriptors- Acoustics, Electronic Equipment, Infant Behavior, *Infants, Language Research, Measurement Techniques, *Oral Expression, *Spectrograms, *Speech, *Statistical Analysis, Verbal Development

Traditional research methods of recording infant verbal behavior, namely, descriptions by a single observer transcribing the utterances of a single infant in a naturalistic setting, have been inadequate to provide data necessary for modern linguistic analyses. The Center for Research on Language and Language Behavior has undertaken to correct this inadequacy. The Center collected permanent, complete, and continuous records of all vocalizations of two infants during their first five months of life. This data was then processed by new electro-acoustic techniques. In processing one hundred and eight 95-second vocal behavior samples, the computer determined (1) the number of utterances, (2) the duration of each utterance, and (3) the mean and standard deviation of the fundamental frequency and amplitude of each utterance. Further statistical analyses are now in progress. (WD)

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Development of the Prosodic Features of Infants' Vocalizing

ED025331

Harlan Lane and William Sheppard

Center for Research on Language and Language Behavior

ED025331

The vocal behavior of an infant during the first few months of life is the matrix for later language development. Therefore, the differentiation and organization of infant vocalizing as a function of maturational and environmental factors holds considerable interest.

Previous research in this area consists almost entirely of descriptions by a single observer transcribing the utterances of a single infant in naturalistic settings. Two abridged biographical reports will be samples: one from a "pre-objective" period, that of the late 19th century, and one from recent decades, employing "more refined techniques" (McCarthy, 1946). The earlier writer, a keen observer of behavior, is Charles Darwin. The contemporary writer, a linguist, is W. F. Leopold.

"The noise of crying is uttered in an instinctive manner... After a time the sound differs as to the cause, such as hunger and pain...he soon appeared to cry voluntarily... When 46 days old he first made little noises without any meaning to please himself, and these soon became varied....At exactly the age of a year, he made the great step of inventing a word for food, namely mum...and now, instead of beginning to cry when he was hungry, he used this word in a demonstrative manner...implying 'Give me food'" (Darwin, 1877, p. 285).

"During the first few weeks the only sounds produced were cries of dissatisfaction... in the seventh and eighth weeks the sounds ceased to be purely incidental. She uttered more arbitrary sounds of satisfaction...cooing as an articulated expression of feelings of satisfaction was therefore well established by the end of the second month... By the seventh month there was a good deal of babbling prevalently ranging from [a] to [ɛ], long, without many tongue movements... At the end of the eleventh month, her active vocabulary consisted of two words" (Leopold, 1939, p. 72).

Linguistic studies in this field have focused on providing a description of the language of a particular child at different levels of development. The units of analysis have been phonemes, morphemes, words and sentences. The procedures for data collection usually begin with phonetic transcription of the infant's vocal behavior by a trained observer. The most

PS001492

extensive studies of this type are those of the linguists Gregoire, Leopold, Cohen, and Velten. The work of Irwin and Chen in the 1940's is notable for the refinements they introduced. These investigators began the practice of using two observers rather than one; this allowed them to obtain measures of observer agreement. They also increased the numbers of subjects observed and selected a more adequate sample of subjects.

Psychologists working in this area have tended to study more molar aspects of language development. The main interest has been to provide normative data on various indices of development, e.g., use of form class, size of vocabulary, mean sentence length, and the like. The procedure for data collection is again transcription (usually alphabetic) by a trained observer using cross-sectional and longitudinal sampling. Among the classic studies of this type are those of Davis, Fisher, McCarthy, Shirley, Templin, and Lewis.

The most striking feature of the research literature on infant vocalizing is the lack of advancement in the field. The basic drawback is the reliance on transcriptions obtained by observations in naturalistic settings. Consider the difficulty of transcribing infant speech sounds. Because the infant is in the process of learning to articulate, the sounds he utters are unlikely to fit neatly into any classificatory system. There is the danger that the trained observer filters the variagated vocal behavior through his own classificatory categories--categories developed with adult vocal behavior--and thus rejects much of importance. Moreover, the infant utters sounds rapidly and sporadically, making it difficult for the observer to keep an accurate or complete record.

The obvious sampling problems have also limited the generality of the findings of previous studies. It has been recognized that when investigator and subject are also mother and child experimental rigor receives little nourishment. In those less numerous studies in which the investigator has intruded into the home, the schedule of observation and transcription usually has been to be most charitable, unsystematic. In summing up the results reported up to 1941 Irwin says:

"It will be apparent from this review of the more important studies...that there does not exist a large body of data secured from adequate samplings of infants for purposes of a statistical analysis... Usually no systematic research methods were formulated; statistical

techniques essential to the analysis of mass data are practically absent, no reliabilities of observers have been established, many observers used alphabetical rather than phonetic systems of symbols for recording; and most reports indulge in an inordinate amount of interpretation supported by very little empirical material." (Irwin, 1941, p. 285).

In her classic review of the literature on language development in the child, McCarthy writes: "Although this wealth of observational material has proven stimulating and suggestive for later research workers, it has little scientific merit, for each of the studies has employed a different method, the observations have for the most part been conducted on single children who were usually either precocious or markedly retarded in their language development; the records have been made under varying conditions, and most of the studies are subject to the unreliability of parents' reports" (McCarthy, 1946, p. 478).

Attempts to ameliorate one or more of the methodological problems that we have reviewed have waited upon advancements in instrumentation. In Part II of her 1929 review McCarthy described some of the early attempts to record speech by means other than transcription by an observer. The first device she described was the manometric flame, invented by Koenig in 1862. With this device, the flame of a gas jet is disturbed by the sound wave and these disturbances are recorded photographically. The phonautograph, invented by Scott in 1859, recorded the speech wave by means of a stylus attached to a diaphragm. The best device of this general type was the phonodeik, which consisted of a horn and a diaphragm; a small platinum wire attached at one end to the diaphragm was passed around a jewel-mounted spindle to a delicate spring. Attached to the spindle was a tiny mirror which reflected a fine beam of light onto a moving photographic film. The phonodeik could respond up to 10,000 cps but the horn and diaphragm introduced a certain amount of distortion.

In 1893 Blondel had devised the oscillograph and with improvements in amplifiers it promised to be a very useful piece of apparatus. However, even before the oscillograph had been developed to the point where it gave a practically perfect representation of the sound wave, it became clear that, once recorded in all its complexity, the waveform was virtually impossible to analyze in ways that were useful for studying speech.

Sound recording devices, the phonograph and later the wire and tape recorders, even when developed further than they were in 1929, solved only part of the problem. The actual speech sounds could then be recorded and replayed but an observer still had to rely on his judgment as a perceiver of sound and speech in order to analyze the data.

Two years after the introduction of the sound spectrograph, in 1949, Lynip published a report of the use of this device in the study of infant speech. He recorded the speech sounds of a little girl, beginning with the birth cry and sampling at intervals ending at 56 weeks when intelligible speech began. A spectrographic analysis of these recordings revealed nothing which resembled the spectrograms of phonemes produced by adults.

In 1960, Winitz published a data-garnished polemic on the subject of the spectrographic analysis of infant vocalization. He argued that the fact that the infant's spectrograms didn't look like spectrograms of adult speech simply proved that the spectrograph isn't a good device for the study of infant vocalizations. "The basic data against which any instrumental method of phonetic analysis must be validated are the phonetic analyses made by competent observers whose validity Lynip questions" (Winitz, 1960, p. 173).

Without resolving the question of the validity of spectrographic description, it must be acknowledged that for each two-second sample of vocalization approximately fifteen minutes are required to process, calibrate, crudely quantify, and classify each spectrogram--and that the problems of observer interpretation remain, although transferred from auditory to visual modes.

This brief account of the techniques that have been employed so far for collection and acoustic analysis of infant vocal behavior indicates that an extension of our knowledge of vocal development requires new techniques. Accordingly, the Center for Research on Language and Language Behavior undertook a year ago to collect permanent, complete and continuous records of all vocalizations of two infants, and then to process these records by novel electro-acoustic techniques.

Recording and sampling of vocalizing. During the deliveries of the infants whose vocalizing is the subject of this study, medical personnel wore lapel microphones whose outputs were recorded on magnetic tape. All subsequent vocalizing by the infants at the

hospital was recorded by placing them in private rooms containing a microphone wired to a fast-acting voice-operated switch (Miratel) and to a tape recorder (Tandberg). After leaving the hospital, both children were cared for at home in plexiglass "air cribs" (T.M.I.) that provided no sources of sound within the crib and attenuation of external sounds--hence, a good recording environment. The parents of the children (research assistants at the Center in both cases) were paid to keep a detailed record on prepared forms of major environmental events affecting the infant. These records were synchronized with the tape recordings by writing down the reading of the footage indicator on the tape recorder.

Complete recordings of all vocal behavior during the first five months of life constitute a formidable tape library, which was sampled for analysis in the following way. A master tape was prepared for each child which contained three 95-sec samples of the vocal behavior during every fourth day of life for the first 141 days. For the samples taken from the first month (in which the infant had no regular sleeping times) the three daily samples were excerpted from the recordings for 12 a.m. to 8 a.m., 8 a.m. to 4 p.m., and 4 p.m. to 12 p.m., respectively. This was accomplished by listening to the recording, beginning at the start of each period, and copying the first 95 sec onto the master tape. In some cases undesired noises intruded and the first 95 sec excluding these intrusions was copied. For the samples of vocalizing in the following three months, the three daily periods from which 95-sec samples were taken were: time at awakening (T) to T + 4 hours; T + 4 hours to T + 8 hours and T + 8 hours to T + 12 hours. These sampling procedures yielded 108 95-sec samples for the initial acoustic analysis.

Analysis of the prosodic features. The development of the prosodic features of the infant's vocal behavior was analyzed by extracting three acoustic parameters of the vocalizing during each of the 108 samples, using analog electronic devices. The outputs of these parameter extractors were sampled every 25 sec by an analog-to-digital converter, then processed by an on-line digital computer (PDP-4, Digital Equipment Corp.).

The changing fundamental frequency of the vocalizing was extracted by filtering tape-recorded signals into two frequency ranges. Since the harmonics of the fundamental frequency often have more energy than the fundamental itself, a range-control voltage is generated

PS001192

when there is energy in the lower range which turns off the upper range to exclude the harmonics. If no energy exists in the lower range, however, the fundamental frequency in the upper range is processed unimpeded. In either case, the nearly sinusoidal output from the filters is amplified in a mixer and read on a frequency meter which provides a DC voltage output proportional to the frequency of the fundamental sine wave at its input. A DC amplifier then adjusts the voltage range and polarity for input to the computer.

The changing amplitude envelope of the vocalizing was extracted by applying the recorded signals to a full-wave rectifier followed by a low-pass filter. The output of this device is a DC voltage that is proportional to the absolute value of the amplitude of the vocal waveform (integrated over approximately one period of the fundamental).

The duration of each utterance within a sample, the third prosodic parameter, was determined in the computer by processing the input from the amplitude extractor. When the amplitude dropped below a threshold value and remained there longer than the silence threshold ($\underline{t}_o$) four out of five samples, the end of an utterance was logged at the time of the initial drop. The start of a new utterance was recorded when the amplitude exceeded threshold again.

In addition to defining the beginning and end of utterances, the computer performed the following preliminary processing. Whenever the amplitude fell below a minimal threshold value $\underline{a}_o$, or the frequency fell below a minimal value $\underline{f}_o$, in a 25 msec sample, the values of $\underline{a}$ and $\underline{f}$ were set to zero. This eliminated spuriously low readings due to noise as well as vocal sounds without voicing at the glottis and hence without prosodic value. It also eliminated false frequency readings that would result from the rise-decay time of the frequency meter in response to instantaneous onset or cessation of voicing.

After sampling and then correcting the amplitude and frequency inputs in this fashion, the digital values were reconverted to voltages and plotted as a function of time on a strip-chart recorder. These records of the amplitude and frequency contours after preliminary processing were compared with those obtained directly from the parameter extractors (before computer processing) so as to choose values of a_o, f_o, and t_o that did not distort the original records.

After the preliminary processing, the computer determined, for each 95-sec sample, the number of utterances as defined above, the duration of each utterance; and the mean and standard deviation of the fundamental frequency and amplitude of each utterance. Pooling these statistics for each of the utterances in a sample, the computer determined next their frequency distributions over the entire sample. These frequency distributions were found to be highly right-skewed. A logarithmic transformation was then applied in order to eliminate the skewness and thus to normalize the distributions. Hence, all statistics were computed using the logarithms of the frequency, amplitude or duration values. The computer determined next the means and standard deviations associated with these transformed composite distributions. Consequently, there were two kinds of statistics reported for each 95-sec sample: (1) within utterance measures of central tendency and variability, averaged over utterances, and (2) between utterance measures of central tendency and variability. All in all, these composite statistics for each sample were printed out along with their frequency distributions:

(let M = mean, S = standard deviation, f = fundamental frequency, a = amplitude, d = duration.)

M(MF) - overall fundamental frequency

S(MF) - variability between utterances in fundamental frequency

M(Sf) - overall variability within utterances in fundamental frequency

S(Sf) - variability between utterances in the variability within utterances in fundamental frequency

M(Ma), S(Ma), M(Sa), S(Sa) - as above but for the amplitude parameter

Md - mean utterance duration

Sd - variability in utterance duration

These statistics, describing the three prosodic features of the vocalizing in each sample, are then plotted separately as a function of age at time of the sample, with the time of day (in three intervals) as a parameter. In this way, developmental trends may be discerned in the prosodic features of the infant's vocal behavior.

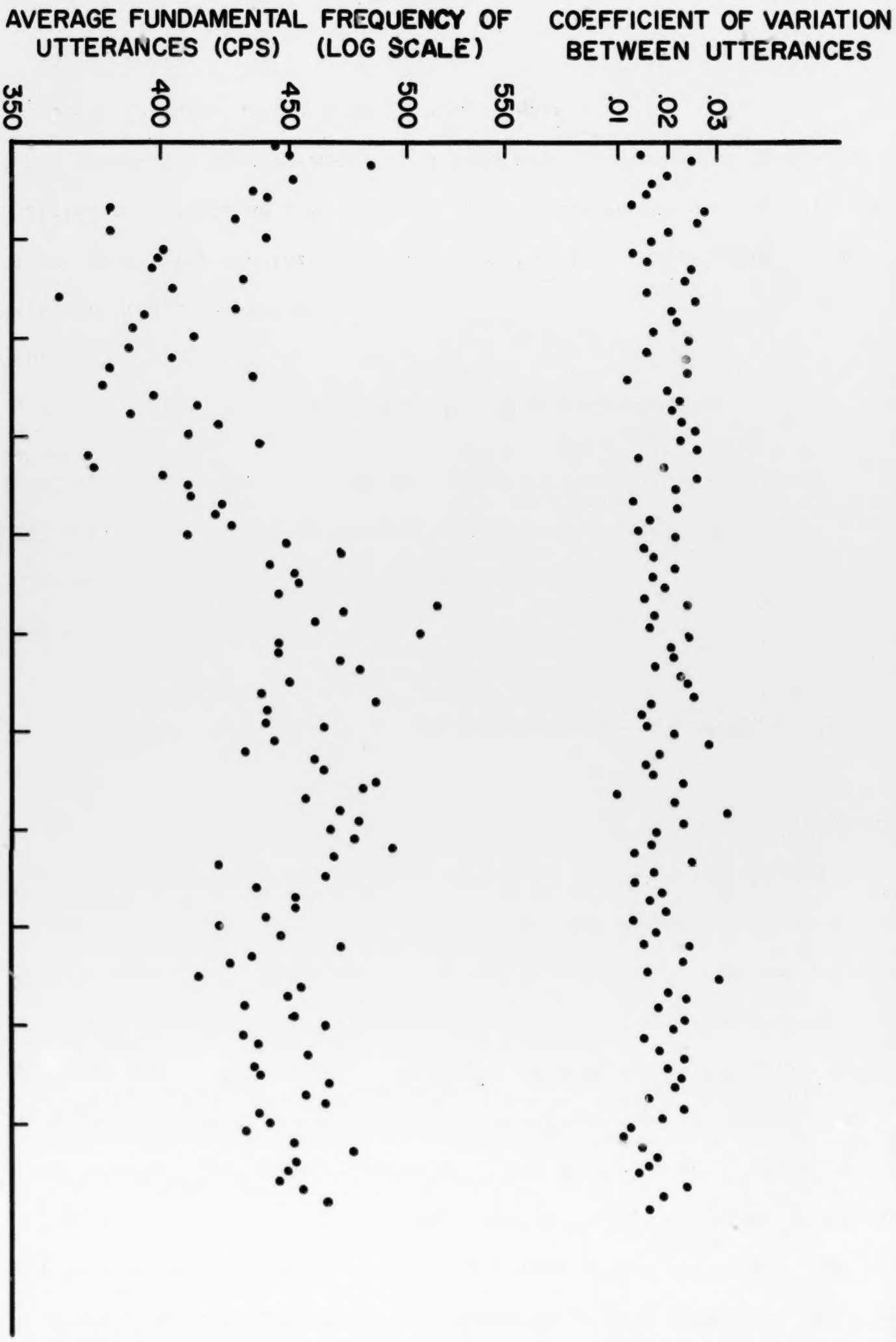
AVERAGE FUNDAMENTAL FREQUENCY OF UTTERANCES (CPS) (LOG SCALE)
350
400
450
500
550
COEFFICIENT OF VARIATION BETWEEN UTTERANCES
.01
.02
.03
10
20
30
40
50
60
70
80
90
100
SAMPLE NUMBER

FIGURE 1

Results. In Fig. 1 the average fundamental frequency of utterances, M(Mf), and the coefficient of variation between utterances in fundamental frequency, CVF_{bU}, [=S(Mf)/M(Mf)] are presented as a function of sample number. An examination of the developmental changes over the first 108 samples (141 days) shows that the average fundamental frequency M(Mf) at birth was approximately 450 cps, that it decreased to 370 cps by sample number 33 (approximately 45 days), and that it then rose and stabilized at about 450 cps for the duration of the study. The coefficient of variation between utterances remained small and constant at between .01 and .03 over the entire study.

Insert Fig. 1 about here

In Fig. 2 the average duration of utterances in msec and the coefficient of variation of duration (CVD_{bU} = Sd/Md) are presented as a function of sample number. The average duration ranges from 100 msec to 800 msec over the 108 samples. No developmental trend is apparent although the variability from sample to sample does decrease with age. The coefficient of variation of duration remains constant at about .20 over the entire sample.

Insert Fig. 2 about here

Further examination of developmental trends awaits the completion of further statistical computations now in progress with the present data.

Footnote

[1]This paper was presented at the First Symposium of the Development of Language Functions, sponsored by the Center for Human Growth and Development, University of Michigan, on October 20-22, 1965.

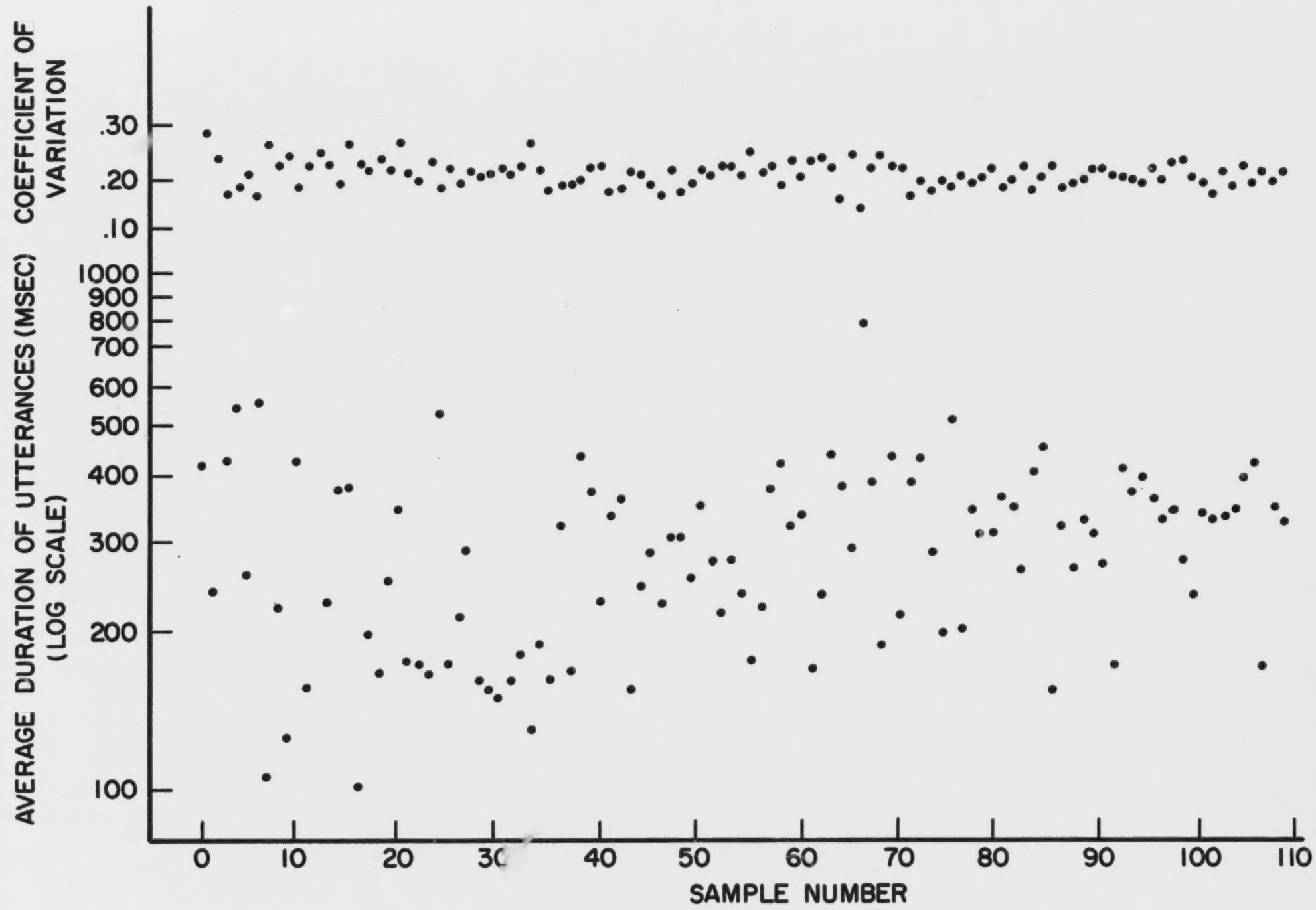
COEFFICIENT OF VARIATION
.30
.20
.10
AVERAGE DURATION OF UTTERANCES (MSEC) (LOG SCALE)
1000
900
800
700
600
500
400
300
200
100
0
10
20
30
40
50
60
70
80
90
100
110
SAMPLE NUMBER

FIGURE 2

References

Carroll, J. B. Language development in children. In S. Saporta (Ed.), Psycholinguistics, New York, Holt, Rinehart and Winston, 1961, Pp. 331-342.

Darwin, C. Biographical sketch of the infant. Mind, 1877, 2, 285-294.

Davis, E. A. Developmental changes in the distribution of parts of speech. Child Developm., 1938, 9, 309-317.

Fisher, M. S. Language patterns of preschool children. Child Developm. Monogr., 1934, No. 15.

Irwin, O. C. Development of speech during infancy: curve of phonemic frequencies. J. exp. Psychol., 1947, 37, 187-193.

Irwin, O. C. Phonetical description of speech development in childhood. In L. Kaiser (Ed.), Manual of Phonetics, Amsterdam, North Holland Publ., 1951, Pp. 403-425.

Irwin, O. C. Research on speech sounds for the first six months of life. Psychol. Bull., 1941, 38, 277-285.

Irwin, O. C. Speech in the young child. J. Speech Hear. Disord., 1952, 17, 269-279.

Irwin, O. C. and Chen, H. P. Development of speech during infancy: curve of phonemic types. J. exp. Psychol., 1946, 36, 431-436.

Irwin, O. C. and Curry, T. Vowel elements in the vocalizations of infants under ten days of age. Child Developm., 1941, 12, 99-109.

Leopold, W. F. Speech development of a bilingual child. Evanston, Ill., Northwestern Univ. Press, 1939.

Leopold, W. F. Bibliography of child language. Evanston, Ill., Northwestern Univ. Press, 1952.

Lewis, M. M. Infant speech: a study of the beginnings of language. New York, Humanities Press, 1951.

Lynip, A. W. The use of magnetic devices in the collection and analysis of the perverbal utterances of the infant. Genet. psychol. Monogr., 1951, 44, 221-262.

McCarthy, D. Language development in children. In L. Carmichael (Ed.), A manual of child psychology, New York, Wiley, 1946, Pp. 476-581.

McCarthy, D. The vocalizations of infants. Psychol. Bull., 1929, 26, 625-651.

Shirley, M. M. The first two years: a study of 25 babies. Intellectual Development, Institute of Child Welfare Monograph, Series No. 7, Minneapolis, Univ. Minnesota Press, 1933.

Skinner, B. F. Verbal behavior. New York, Appleton-Century-Crofts, 1957.

Velten, H. V. The growth of phonemic and lexical patterns in infant language. Language, 1943, 19, 281-292.

Winitz, H. Spectrographic investigation of infant vowels. J. genet. Psychol., 1960, 96, 171-181.